

Mapping Genes Encoding Microsomal ω -6 Desaturase Enzymes and Their Cosegregation with QTL Affecting Oleate Content in Soybean

Eleni Bachlava,* Ralph E. Dewey, Jérôme Auclair, Sanbao Wang, Joseph W. Burton, and Andrea J. Cardinal

ABSTRACT

The microsomal ω -6 desaturase enzymes, which catalyze the desaturation of oleic acid to linoleic acid during fatty acid biosynthesis, are encoded by the *FAD2-1* and *FAD2-2* genes in soybean [*Glycine max* (L.) Merr.]. Breeders aim to incorporate the high-oleate trait into soybean germplasm in order to improve the nutritional value and oxidative stability of soybean oil. The objectives of this study were to map the isoforms of the *FAD2-1* and *FAD2-2* genes and investigate the association of these genetic loci with the oleate phenotype in three populations segregating for oleate content. The populations were grown in replicated multienvironment field trials. According to linkage analysis conducted for two of the populations, *FAD2-1A* and *FAD2-1B* mapped on Linkage Groups O and I, respectively, while the closely linked *FAD2-2A* and *FAD2-2B* isoforms mapped on Linkage Group L. Oleate quantitative trait loci with minor effects were detected in the proximity of *FAD2-1B* and possibly *FAD2-2B* on Linkage Groups I and L. Quantitative trait loci affecting maturity were also detected on chromosomal regions adjacent to the *FAD2* genes. The genotyping assays developed for the *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* isoforms, as well as their linked simple sequence repeat markers, can be used in soybean breeding programs for the elevation of oleic acid seed content through marker-assisted selection.

E. Bachlava, R.E. Dewey, J. Auclair, S. Wang, and A.J. Cardinal, Dep. of Crop Science, North Carolina State Univ., Campus Box 7620, Raleigh, NC 27695-7620; J.W. Burton, USDA-ARS, Soybean and Nitrogen Fixation Research Unit, 3127 Ligon St., Raleigh, NC 27607. Received 14 Oct. 2007. *Corresponding author (ebachla@ncsu.edu).

Abbreviations: ASPE, allele-specific primer extension; BLUP, best linear unbiased predictor; bp, base pair; CAPS, cleaved amplified polymorphic sequence; FAD, fatty acid desaturase; QTL, quantitative trait locus; SSR, simple sequence repeat; SBE, single-base extension; SNP, single nucleotide polymorphism.

SOYBEAN [*Glycine max* (L.) Merr.] is an oilseed crop of major importance, providing much of the world's edible oil. According to the National Agricultural Statistics Service (www.nass.usda.gov/; verified 18 Jan. 2008), soybean oilseeds produce 29.4% of the world's edible vegetable oil, while the United States represent 40.1% of the world's soybean production. Yet, soybean oil is lacking in some quality characteristics. Oil quality is determined by the relative composition of saturated and unsaturated fatty acids in seed triacylglycerols. The average fatty acid composition of soybean oil is 110 g kg⁻¹ palmitic acid (16:0), 40 g kg⁻¹ stearic acid (18:0), 230 g kg⁻¹ oleic acid (18:1), 530 g kg⁻¹ linoleic acid (18:2), and 90 g kg⁻¹ linolenic acid (18:3) (Wilson, 2004). Increased concentration of the monounsaturated oleic acid would improve the nutritional value and oxidative stability of soybean oil. This would also reduce the need for hydrogenation that leads to undesirable *trans* fatty acids (Wilson, 2004). Thus, extensive research has been conducted to incorporate the high-oleic-acid trait into soybean germplasm using traditional plant breeding methods (Wilson et al., 1981; Burton et al., 1989, 2006)

Published in Crop Sci. 48:640–650 (2008).

doi: 10.2135/cropsci2007.07.0381

© Crop Science Society of America

677 S. Segoe Rd., Madison, WI 53711 USA

All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Permission for printing and for reprinting the material contained herein has been obtained by the publisher.

and genetic engineering approaches (Kinney, 1995; Kinney and Knowlton, 1998; Buhr et al., 2002).

Fatty acids are the precursors for the synthesis of triacylglycerols (TAGs), which constitute the storage lipid reserves of the developing oilseeds. Polyunsaturated fatty acids of the seed TAGs are synthesized primarily by ω -3 and ω -6 desaturase enzymes located in the endoplasmic reticulum; however, a small percentage is also derived from comparable enzymes localized in the plastids. The desaturase enzymes of the endoplasmic reticulum act on the fatty acids esterified to the first and second backbone position of phosphatidylcholine (PC), whereas the plastid-localized desaturases predominantly utilize galactolipids as substrates (Ohlrogge and Browse, 1995).

Plant fatty acid desaturase genes were initially elucidated in *Arabidopsis*. The desaturation of oleic to linoleic acid in the endoplasmic reticulum was shown to be controlled by the *fad2* gene, encoding the oleoyl-PC ω -6 desaturase activity (Miquel and Browse, 1992; Okuley et al., 1994). Subsequently, genes designated *FAD2-1* and *FAD2-2* were reported to encode comparable activities in soybean (Heppard et al., 1996). The *FAD2-1* gene was expressed only during seed development and *FAD2-2* was constitutively expressed during seed development and vegetative growth. Therefore, it was suggested that the seed-specific *FAD2-1* is primarily responsible for the desaturation of oleic to linoleic acid for fatty acids destined for seed storage lipids, while *FAD2-2* is involved in the desaturation of membrane lipids in vegetative tissues, playing a lesser role in storage oil biosynthesis. Moreover, the existence of at least two copies of each of the *FAD2-1* and *FAD2-2* genes was proposed (Heppard et al., 1996).

Recently, the existence of two distinct isoforms of the soybean *FAD2-1* gene (designated *FAD2-1A* and *FAD2-1B*), encoding microsomal ω -6 desaturase enzymes that differ in stability at high temperatures, was confirmed (Tang et al., 2005). The *FAD2-1A* and *FAD2-1B* genes map on Linkage Groups O and I, respectively, in the soybean genome. Both *FAD2-1A* and *FAD2-1B* produce alternatively spliced transcripts that do not change the amino acid residues of the encoded ω -6 desaturase enzymes (Schlueter et al., 2007). For *FAD2-2* genes, which also encode microsomal ω -6 desaturase enzymes, three isoforms have been reported, designated *FAD2-2A*, *FAD2-2B*, and *FAD2-2C*. *FAD2-2A* and *FAD2-2B* are localized on Linkage Group L (Schlueter et al., 2007).

Environmental factors affect the unsaturated fatty acid content of soybean oilseeds. Temperature exerts an effect on the unsaturated fatty acid composition during the pod-filling stage that coincides with oil deposition. Planting and maturity dates lead to changes in the fatty acid profiles due to different temperatures during oil deposition (Burton et al., 1983; Wilcox and Cavins, 1992). For example, the elevated-oleic lines N98-4445A and N97-3363-3, which

were used in this study, are unstable across environments and their oleate variation is largely attributed to growth temperature differences (Oliva et al., 2006; Wilson et al., 2002). Although temperature effects during oil deposition were implicated in the altered fatty acid composition of soybean oilseeds, the expression patterns of the isoforms encoding the ω -6 desaturase enzymes, with the exception of *FAD2-2C*, showed no change in the level of transcript accumulation at different temperatures during seed development that could explain the changes in the levels of unsaturated fatty acids in seed triacylglycerols (Heppard et al., 1996; Schlueter et al., 2007). A recent study of the ω -6 desaturase enzymes encoded by the *FAD2-1A* and *FAD2-1B* genes, however, provided evidence of differential enzymatic stability at high temperatures (Tang et al., 2005). Apart from temperature effects, coordination of metabolite flux through the reactions of glycerolipid biosynthesis in the endoplasmic reticulum coupled with the de novo fatty acid biosynthesis pathway in the plastids underlie the complexity of the factors (both genetic and environmental) that can ultimately determine oleate composition of the storage oil.

Although the location of the isoforms encoding the ω -6 desaturase enzymes was recently revealed, there is minimal information regarding whether the isoforms of *FAD2-1* and *FAD2-2* genes cause the observed genotypic variation for oleate content in populations in which this trait is segregating. As an initial step toward unraveling the complexity of oleate biosynthesis and accumulation in soybean, the effects of the *FAD2-1* and *FAD2-2* isoforms (and their alleles), as well as their epistatic and environmental interactions, need to be investigated (Cardinal, 2008). The objectives of this study were to: (i) identify single nucleotide polymorphisms (SNPs) in the coding or upstream regions of the isoforms of *FAD2-1* and *FAD2-2* genes for the soybean lines under study; (ii) develop isoform- and allele-specific markers; (iii) map the individual isoforms of the *FAD2-1* and *FAD2-2* genes in the soybean genome; and (iv) investigate whether segregation at *FAD2-1* and *FAD2-2* loci is associated with changes in oleic acid seed content in three soybean populations segregating for oleate content.

MATERIALS AND METHODS

Development of Soybean Populations

Three soybean populations, designated as FAE, FAF, and FAS, were developed by single-seed descent (Brim, 1966). Population FAE consists of 721 F_5 -derived lines from the cross of N98-4445A \times N97-3525. Population FAF represents 118 F_5 -derived lines from the cross of N97-3363-3 \times PI423893. Population FAS consists of 231 F_5 -derived lines from the cross of N98-4445A \times PI423893. The N98-4445A line is a high oleic (563.1 g kg⁻¹), low linoleic acid line developed by the USDA-ARS in Raleigh, NC (Burton et al., 2006). The N98-4445A line originated as a plant selection from the cross N94-2473 \times (N9

3-2007-4 × N92-3907) and is a sister line of N97-3363-3. The PI423893 line is a plant introduction with mid-oleic acid (305.8 g kg⁻¹) seed oil content and unknown genetic background (USDA-ARS National Plant Germplasm System, www.ars-grin.gov/npgs/searchgrin.html, verified 12 Jan. 2008). The N97-3525 line is the low palmitate, low linolenate cultivar Satellite, which was derived from the cross of 'Soyola' × {'Brim' (2) × [N88-431(2) × (N90-2013 × C1726)]} (Cardinal et al., 2007). The FAF and FAS populations are segregating for oleate genes and the reduced-linolenate *fan*(PI123440) allele, while the FAE population is segregating for oleate genes and the *fan_{nc}* reduced-palmitate allele.

Experimental Design

The FAE and FAF populations were planted separately in sets-within-replications experimental designs with two replications in each location. Fifteen sets of seven rows by seven columns were randomly assigned in each replication, location, and year for the FAE population. Five sets of five rows by five columns were randomly assigned in each replication, location, and year for the FAF population. In each population, before lines were randomly assigned to each set, they were divided into five arbitrary groups that differed in maturity according to the lines' maturity phenotypes in 2004. Each maturity group was represented at least once in each set of the FAE and FAF populations, and the maturity groups that included the majority of the segregating lines were represented two to three times in each set. Once lines were randomly assigned to a set, they remained in that set in all replications, locations, and years. Within each set, the maturity groups were randomly assigned to a column to facilitate mechanical harvest. Then the lines within each maturity group were randomly assigned to plots within each column. The parental lines N97-3363-3 and PI423893 in the FAF population and the parental line N97-3525 in the FAE population were randomly assigned to one set, to a maturity group within that set depending on their maturity date, and, finally, to a plot within the set, as explained above. The N98-4445A parental line of the FAE population was assigned to a maturity group in all the sets and randomly assigned a plot within each set as explained above. The FAE population was planted in four-row plots at Clinton, Kinston, and Plymouth, NC, in 2005 and 2006. The FAF population was grown at Clinton and Kinston in 2005 and at Clinton, Kinston, Clayton, and Plymouth, NC, in 2006. The FAF population was planted in four-row plots, with the exception of Clayton in 2006, which was planted in one-row plots.

The FAS population was planted in five-seed hills at Clayton, NC, in a sets-within-replications design with three replications in 2006. Sets were randomly assigned to each replication, and lines within each set were randomly assigned to hills. The parental lines, N98-4445A and PI423893, were randomly assigned within each of the five sets in each replication.

Phenotypic Evaluation

Maturity date was recorded for the FAE, FAF, and FAS populations at the R8 reproductive stage as days after planting (Fehr and Caviness, 1977). For the FAE and FAF populations, mature soybean seed was harvested mechanically from the two middle rows of each experimental plot or each single-row plot for the FAF population at Clayton in 2006. Approximately 10 g of seed was subsampled from each plot for the evaluation of fatty acid composi-

tion. For the FAS population, 5 g of seed was subsampled from the seed harvested from each hill. Fatty acid composition was evaluated by gas chromatography, as described by Burkey et al. (2007).

Statistical Analysis

Statistical analysis of the sets-within-replications design was conducted using Proc MIXED in SAS 9.1 (SAS Institute, 2004). Environments, sets, replications, and lines, as well as their interactions, were considered as random effects. Best linear unbiased predictors (BLUPs) were obtained for all traits and lines of the FAE, FAF, and FAS populations. The degrees of freedom for FAF and FAS populations were calculated using the Satterthwaite approximation, while for the FAE population they were derived with the containment method to reduce the computing power required. Analysis of the FAE and FAF populations was conducted separately for each environment, as well as combined across all the environments for fatty acid traits and maturity date. The phenotypic data of the FAF population from Clinton in 2005 were discarded due to numerous missing data points and heterogeneity of variance in comparison with the other environments. The analysis of the FAS population for a single environment was performed as described above.

Simple sequence repeat (SSR) markers and allele-specific markers were tested for association with fatty acid phenotypes, derived as BLUPs, in FAE, FAF, and FAS populations using single-factor analysis in Proc GLM of SAS 9.1. The additive genetic effects were obtained by the CONTRAST statement in Proc GLM of SAS 9.1. Dominance genetic effects were also estimated for the FAS population. The means of the homozygous genotypic classes for each trait in the FAE and FAF populations or the homozygous and heterozygous genotypes in the FAS population, as well as their standard errors, were obtained with the LSMEANS statement in Proc GLM of SAS 9.1. The difference of the means of the homozygous genotypic classes, as well as its standard error and its significance, were determined with the ESTIMATE statement in Proc GLM of SAS 9.1. To investigate whether maturity date affected the associations between fatty acid traits and molecular markers, single-factor analysis was also conducted, with maturity date included as a covariate in the model.

Genotypic Evaluation and Linkage Mapping

Genomic DNA was extracted from leaf tissue of approximately 5 to 10 plants for each of the lines of the FAF and FAE populations collected from Kinston (2005) and Clayton (2004), respectively. The DNA isolation was conducted in a 96-well plate format with the Gentra PureGene DNA purification kit (Gentra Systems, Minneapolis, MN). Genomic DNA was isolated from leaf tissue (Keim et al., 1988) of each line of the FAS population from all the plants in each hill grown at Clayton (2006). One hundred forty-seven polymorphic SSR markers of the consensus linkage map (Cregan et al., 1999; Song et al., 2004) were genotyped across the 20 linkage groups of the soybean genome in the FAF population. Only SSR markers flanking the genomic regions of interest, as discussed below, were genotyped in the FAE and FAS populations (Table 1). The reactions were performed in a 384-well PTC-200 Peltier Thermal Cycler (MJ Research, Waltham, MA). Thermocycling conditions were 95°C for 2 min and 38 cycles of 92°C for 1 min, 49°C for 1 min, and 68°C for 1 min and 30 s. The amplifica-

tion products were resolved on 4% super-fine resolution agarose gels (Amresco, Solon, OH) with ethidium bromide staining in 1x Tris-borate-EDTA buffer, or 6.5% polyacrylamide gels in a LICOR 4300 DNA analysis system (LICOR Biosciences, Lincoln, NE) using M13-tailed unmodified primers and IRD-700 or IRD-800 labeled M13 oligonucleotides.

Linkage analysis was performed with JoinMap 3.0 (Van Ooijen and Voorrips, 2001). The Kosambi's mapping function was used for linkage map construction and linkage was declared with a maximum recombination frequency of 0.4 and a minimum logarithm of odds (LOD) score of 3.0. Linkage analysis was also conducted with MAPMAKER/EXP 3.0 (Lander et al., 1987), as previously described by Cardinal et al. (2001), to verify the insufficient linkage declared in JoinMap 3.0 and the differential placement of *FAD2-2B* in FAF and FAS populations, as explained below. Linkage analysis with MAPMAKER/EXP 3.0 was performed using the same LOD and recombination thresholds as linkage analysis with JoinMap 3.0; however, the heterozygote genotypes were discarded in order for the F_5 -derived lines of the FAF population to be analyzed as recombinant inbred lines, while the F_3 -derived lines of the FAS population were actually analyzed as F_2 lines. Therefore, linkage mapping results from both analyses are discussed here, but only the linkage analysis with JoinMap 3.0 is presented.

Database Search and Polymerase Chain Reaction Primer Design

The coding regions of *FAD2-1A* and *FAD2-1B* were initially amplified with isoform-specific primers designed from the GenBank (Benson et al., 2002) accessions AB188250 and AB188251, respectively. For the isoforms of *FAD2-2*, *in silico* analysis was conducted on the soybean expressed sequence tags (Shoemaker et al., 2002) deposited in GenBank using BLASTN searches. Three distinct isoforms were predicted and isoform-specific primers were designed. The specific primers that amplify the coding region of each isoform are: *FAD2-1A_d_F* (5'-GTG TGG CCA AAG TGG AAG TT-3') and *FAD2-1A_b_R* (5'-CAA AGC TCC CTT CAG CCA GT-3'); *FAD2-1B_a_F* (5'-GTG TGG CCA AAG TTG AAA T-3') and *FAD2-1B_b_R* (5'-AAG CAC CCC TCA GCC AAT-3'); and *FAD2-2B_c_F* (5'-CCT CAG TTT AGT CTC AGC CAG ATT-3') and *FAD2-2B_c_R* (5'-TGG GCT TTA TTG CCT TTG TC-3'). The specific primers that amplify the upstream region of the *FAD2-2B* isoform are: *FAD2-2B_UP_1_F* (5'-ACG CAG TGA ATC AAA TGA CAA-3') and *FAD2-2B_UP_2_R* (5'-CTT CTT AAT CTG GCT GAG ACT AAA CT-3'). Isoform and primer designations for the isoforms of *FAD2-1* genes follow the nomenclature of Tang et al. (2005), while the isoforms of *FAD2-2* genes follow the nomenclature of Schlueter et al. (2007).

Sequencing of *FAD2* Genes, Single Nucleotide Polymorphism Detection, and Allele-Specific Marker Development

Genomic DNA was isolated from bulked leaf tissue (Keim et al., 1988) from approximately 10 plants of the N98-4445A, N97-3363-3, and PI423893 lines and Brim. The amplifica-

tion of the various *FAD2-1* and *FAD2-2* genes was performed in a MyCycler thermal cycler (Bio-Rad, Hercules, CA). The polymerase chain reactions (PCRs) were performed using the Expand High Fidelity PCR system (Roche Applied Science, Indianapolis, IN). Each 50- μ L reaction contained 0.8 U polymerase, 20 pmol of each of the forward and reverse primers, 187 μ mol L⁻¹ deoxynucleotide triphosphates (dNTPs), and 100 ng of genomic template DNA. The cycling program included a 5-min step at 94°C, 30 amplification cycles of 30-s denaturation at 94°C, annealing at 55°C for 30 s, and extension at 72°C for 2 min, followed by an extension step for 7 min at 72°C. The amplification products were cloned into pCR 2.1 cloning vectors (Invitrogen, Carlsbad, CA) and sequenced (Iowa State DNA facility, Ames, IA). Pair-wise alignments of the sequences obtained for the isoforms of *FAD2-1* and *FAD2-2* genes were performed for the detection of SNPs among the soybean lines N98-4445A, N97-3363-3, PI423893, and Brim. Single nucleotide polymorphisms leading to restriction site polymorphisms were identified, and cleaved amplified polymorphic sequence (CAPS) markers were developed for *FAD2-1B* and *FAD2-2B* using the *Hpy*-CH4III and *Alu*I endonucleases, respectively.

The amplified fragments of *FAD2-1B* and *FAD2-2B* isoforms for CAPS analysis were obtained with the following

Table 1. Molecular assays designed for the isoforms *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* as well as simple sequence repeat markers mapping on Linkage Groups I, O, and L that were tested for associations with the unsaturated fatty acids' content in the FAF (N97-3363-3 \times PI423893), FAE (N98-4445A \times N97-3525), and FAS (N98-4445A \times PI423893) populations.

Linkage group	FAF	FAE	FAS
Group I			
<i>FAD2-1B</i> CAPS [†]	x		x
satt571	x		
satt496	x		
satt354	x	x	x
sat_268	x	x	x
sat_418	x		
sat_324	x		
sat_299	x		
Group L			
<i>FAD2-2B</i> CAPS	x		x
satt495	x		
satt652	x	x	x
sat_320	x		
satt462	x	x	x
sat_340	x	x	x
satt006	x		
Group O			
<i>FAD2-1A</i> ASPE/SBE [†]	x		x
satt420	x		
satt478	x		
satt153	x	x	x
sat_108	x		x

[†]CAPS, cleaved amplified polymorphic sequence; ASPE, allele specific primer extension; SBE, single-base extension.

primers: FAD2-1B_c_F (5'-ATC CAA AGT TGC ATG GTA CAC-3'), FAD2-1B_b_R (5'-AAG CAC CCC TCA GCC AAT-3'), FAD2-2B_UP_1_F (5'-ACG CAG TGA ATC AAA TGA CAA-3'), and FAD2-2B_UP_1_R (5'-TGC TAT CGT TAC CCA CAT ACG A-3'). Cycling conditions were the same as described above. Amplification of genomic DNA was conducted with a 12- μ L PCR reaction containing 6 pmol of each of the forward and reverse primers, 0.5 U Taq polymerase (New England Biolabs, Ipswich, MA), 10x buffer (10 mmol L⁻¹ Tris-HCl, 50 mmol L⁻¹ KCl, 1.5 mmol L⁻¹ MgCl₂, pH 8.3), 234 μ mol L⁻¹ dNTPs, and 75 ng genomic DNA. Next, 10 μ L of the amplified product was digested in a 22- μ L reaction at 37°C for 2 h. For the digestion of *FAD2-1B* fragments, the reaction contained 2.5 U *Hpy*CH4III, buffer (20 mmol L⁻¹ Tris-acetate, 50 mmol L⁻¹ potassium acetate, 10 mmol L⁻¹ magnesium acetate, 1 mmol L⁻¹ dithiothreitol, pH 7.9) and 5 μ g bovine serum albumin (BSA). The *FAD2-2B* amplification products were digested in reactions containing of 5 U *Alu*I, buffer (10 mmol L⁻¹ Tris-HCl, 50 mmol L⁻¹ NaCl, 10 mmol L⁻¹ MgCl₂, 1 mmol L⁻¹ dithiothreitol, pH 7.9) and 5 μ g BSA. Restriction fragments were resolved in 4% superfine-resolution agarose gels, as described above.

In the FAF population, the *FAD2-1A* isoform was genotyped with the allele-specific primer extension (ASPE) assay (University of California-Davis Genome Center, Davis, CA) in a Luminex platform (Luminex Corp., Austin, TX) using the primers FAD2-1A_ASPE1 (5'-TCA ATT ACT TCA CTT TAA TCC TTT gcc aaa gtg gaa gtt caa g-3') for N97-3363-3 and FAD2-1A_ASPE2 (5'-TCA TTC ATA TAC ATA CCA ATT CAT gcc aaa gtg gaa gtt caa c-3') for PI423893. The primers include the SNP-specific sequence (lowercase letters) coupled to the oligonucleotides associated with the fluorescent microspheres (uppercase letters). In the FAS population, *FAD2-1A* was genotyped with the single-base extension (SBE) assay in a Luminex platform (Luminex Corp., Austin, TX) according to Chen et al. (2000) (data not shown).

RESULTS AND DISCUSSION

Identification of the Isoforms of the *FAD2-2* Gene

A database query indicated that there are three isoforms for the *FAD2-2* gene represented within GenBank accessions AB188253 and AC166091 (data not shown). The AC166091 accession corresponds to a bacterial artificial chromosome (BAC) containing two independent *FAD2-2* genes. Recently, Schlueter et al. (2007) designated the isoform in AB188253 (also found in a BAC corresponding to accession no. AC166742) as *FAD2-2C* and the two isoforms in AC166091 as *FAD2-2A* and *FAD2-2B*. Because *FAD2-2A* and *FAD2-2B* are located on the same BAC, they should map to the same chromosomal region in the soybean genome. Since *FAD2-2A* has an approximately 100-base pair (bp) deletion within the coding region of the enzyme, it probably encodes a nonfunctional enzyme. Therefore, *FAD2-2A* was not considered a candidate gene that could explain differ-

ences in the ω -6 desaturase activity in our populations and was not studied further. Indeed, transcript analysis of *FAD2-2A* recently provided evidence that this isoform may not be expressed (Schlueter et al., 2007). Also, gene-specific primers designed for *FAD2-2C* amplified only part of the isoform's coding region, which did not reveal any SNPs (data not shown) among the genotypes used in this study. Consequently, further information for the *FAD2-2C* isoform could not be obtained.

Single Nucleotide Polymorphism Detection

The isoform-specific primers were used for the amplification of *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* from genomic DNA of lines N98-4445A, N97-3363-3, and PI423893. Amplification products were cloned, sequenced, and aligned for the parental soybean lines of interest and the cultivar Brim, leading to the detection of SNPs in coding or upstream regions of each isoform. One SNP originating from PI423893 results in the change of the 22nd amino acid residue of the predicted protein of *FAD2-1A* from glycine to arginine due to a substitution of a guanine with a cytosine in the first position of the codon. This is one of the four amino acid residues that differ between the *FAD2-1A* and *FAD2-1B* enzymes in the N-terminal region. Interestingly, the location of the SNP in the coding region of *FAD2-1A* coincides with a region responsible for the instability of the enzyme at high temperatures (Tang et al., 2005).

For *FAD2-1B*, a total of five SNPs were detected in the coding region, two of which result in nonsynonymous amino acid changes. The substitution of a cytosine with a thymine in the second position of the 86th codon of *FAD2-1B* causes an amino acid change from serine to phenylalanine in PI423893. Also, the 126th amino acid residue of the *FAD2-1B* enzyme in N98-4445A and N97-3363-3 changes from valine to methionine due to the substitution of a guanine with an adenine at the first position of this codon. Three additional nucleotide polymorphisms that do not change the amino acid residues of the predicted *FAD2-1B* enzyme were also detected in PI423893 (Fig. 1). These are located in the third positions of codons 219, 223, and 228. All nonsynonymous SNPs in *FAD2-1A* and *FAD2-1B* were evaluated for potential implications in protein function using the SIFT program (Ng and Henikoff, 2003). None of the nonsynonymous amino acid changes, however, were predicted to be deleterious.

Amplification and sequencing of the coding regions of *FAD2-2B* revealed no polymorphisms among the soybean lines of interest. Therefore, primers were designed to amplify a region immediately upstream of *FAD2-2B*. One SNP was detected 641 bp upstream of the putative start codon of *FAD2-2B*, corresponding to a polymorphism of a thymine vs. cytosine.

N98-4445A	<i>ATCCAAAGTTGCATGGT</i> 17
N97-3363-3	<i>ATCCAAAGTTGCATGGT</i> 17
PI423893	<i>ATCCAAAGTTGCATGGT</i> 17
Brim	<i>ATCCAAAGTTGCATGGT</i> 17

N98-4445A	<i>ACACCAAGTACCTGAACAACCTCTAGGAAGGGCTGCTTCTCTCTCATCACA</i> 77
N97-3363-3	<i>ACACCAAGTACCTGAACAACCTCTAGGAAGGGCTGCTTCTCTCTCATCACA</i> 77
PI423893	<i>ACACCAAGTACCTGAACAACCTCTAGGAAGGGCTGCTTCTCTCTCATCACA</i> 77
Brim	<i>ACACCAAGTACCTGAACAACCTCTAGGAAGGGCTGCTTCTCTCTCATCACA</i> 77

N98-4445A	TAGGGTGGCCTTTGTATTTAGCCTTCAATGTCTCTGGCAGACCCATGATGGTTTTGCTA 137
N97-3363-3	TAGGGTGGCCTTTGTATTTAGCCTTCAATGTCTCTGGCAGACCCATGATGGTTTTGCTA 137
PI423893	TAGGGTGGCCTTTGTATTTAGCCTTCAATGTCTCTGGCAGACCCATGATGGTTTTGCTA 137
Brim	TAGGGTGGCCTTTGTATTTAGCCTTCAATGTCTCTGGCAGACCCATGATGGTTTTGCTA 137

N98-4445A	GCCACTACCACCTTATGCTCCCATATATTCAAATCGTGAGAGGCTTTTGATCTATGTCT 197
N97-3363-3	GCCACTACCACCTTATGCTCCCATATATTCAAATCGTGAGAGGCTTTTGATCTATGTCT 197
PI423893	GCCACTACCACCTTATGCTCTTATATATTCAA <u>ACCGT</u> GAGAGGCTTTTGATCTATGTCT 197
Brim	GCCACTACCACCTTATGCTCCCATATATTCAAATCGTGAGAGGCTTTTGATCTATGTCT 197

N98-4445A	CTGATGTTGCTTTGTTTTCTGTGACTTACTTGCTCT <u>ACCGT</u> GTTGCAACTATGAAAGGGT 257
N97-3363-3	CTGATGTTGCTTTGTTTTCTGTGACTTACTTGCTCT <u>ACCGT</u> GTTGCAACTATGAAAGGGT 257
PI423893	CTGATGTTGCTTTGTTTTCTGTGACTTACTTGCTCT <u>ACCGT</u> GTTGCAACTATGAAAGGGT 257
Brim	CTGATGTTGCTTTGTTTTCTGTGACTTACTTGCTCT <u>ACCGT</u> GTTGCAACTATGAAAGGGT 257

N98-4445A	TGGTTTGGCTGCTATGTGTTTATGGGGTGCCATTGCTCATTGTGA <u>ACCGT</u> TTTCTTTGTGA 317
N97-3363-3	TGGTTTGGCTGCTATGTGTTTATGGGGTGCCATTGCTCATTGTGA <u>ACCGT</u> TTTCTTTGTGA 317
PI423893	TGGTTTGGCTGCTATGTGTTTATGGGGTGCCATTGCTCATTGTGA <u>ACCGT</u> TTTCTTTGTGA 317
Brim	TGGTTTGGCTGCTATGTGTTTATGGGGTGCCATTGCTCATTGTGA <u>ACCGT</u> TTTCTTTGTGA 317

N98-4445A	CCATCACATATCTGCAGCACACACTATGCCTTGCCCTCACTATGATTCATCAGAATGGG 377
N97-3363-3	CCATCACATATCTGCAGCACACACTATGCCTTGCCCTCACTATGATTCATCAGAATGGG 377
PI423893	CCATCACATATCTGCAGCACACACTATGCCTTGCCCTCACTATGATTCATCAGAATGGG 377
Brim	CCATCACATATCTGCAGCACACACTATGCCTTGCCCTCACTATGATTCATCAGAATGGG 377

N98-4445A	<i>ATTGGCTGAGGGGTGCTT</i> 395
N97-3363-3	<i>ATTGGCTGAGGGGTGCTT</i> 395
PI423893	<i>ATTGGCTGAGGGGTGCTT</i> 395
Brim	<i>ATTGGCTGAGGGGTGCTT</i> 395

Figure 1. Alignment of the 395-bp fragment amplified for *FAD2-1B* cleaved amplified polymorphic sequence analysis in N98-4445A, N97-3363-3, PI423893, and 'Brim'. The primers *FAD2-1B_c_F* and *FAD2-1B_b_R* designed for amplification are italicized. The restriction sites for *Hpy*CH4III are underlined and the three single nucleotide polymorphisms (SNPs) are highlighted. The substitution from thymine to cytosine in one of the SNPs generates an additional restriction site for *Hpy*CH4III endonuclease in PI423893.

Development of Allele-Specific Molecular Markers for *FAD2-1B* and *FAD2-2B*

The detected nucleotide substitutions were examined for possible restriction site polymorphisms coinciding with the location of the SNPs to develop molecular assays that could readily distinguish among the alleles found in N98-4445A, N97-3363-3, and PI423893. The substitution of a thymine with a cytosine in the 223rd codon of *FAD2-1B* generates an additional restriction site for the *Hpy*CH4III endonuclease in PI423893 (Position 172, Fig. 1), and thus enabled the development of a CAPS marker. Digestion of a 395-bp *FAD2-1B*-specific fragment from PI423893 generates fragments of 172, 90, 70, and 63 bp, whereas the same product derived from N98-4445A or N97-3363-3 CAPS yields fragments of 235, 90, and 70 bp (data not shown). The additional *Hpy*CH4III restriction sites that do not correspond to nucleotide substitutions within the amplified fragment serve as digestion controls.

The unique SNP in the upstream region of *FAD2-2B* results in the loss of an *AluI* restriction site in the fragment amplified from PI423893. The primers used for CAPS analysis amplify a 494-bp fragment in the upstream region of *FAD2-2B*. The *AluI* digestion of this amplification product from N98-4445A and N97-3363-3 generates fragments of size 173, 147, and 174 bp, in contrast to fragments of 320 and 174 bp when genomic DNA from PI423893 is used as a template (data not shown).

The SNP identified in the coding region of *FAD2-1A* did not alter restriction endonuclease cleavage sites. Therefore, an ASPE assay was designed with SNP-specific primers to distinguish between the N97-3363-3 and PI423893 *FAD2-1A* alleles in the FAF population. In the FAS population, an SBE assay was designed according to Chen et al. (2000) to distinguish between the N98-4445A and PI423893 *FAD2-1A* alleles using fluorescent microspheres in a Luminex platform. The segregation

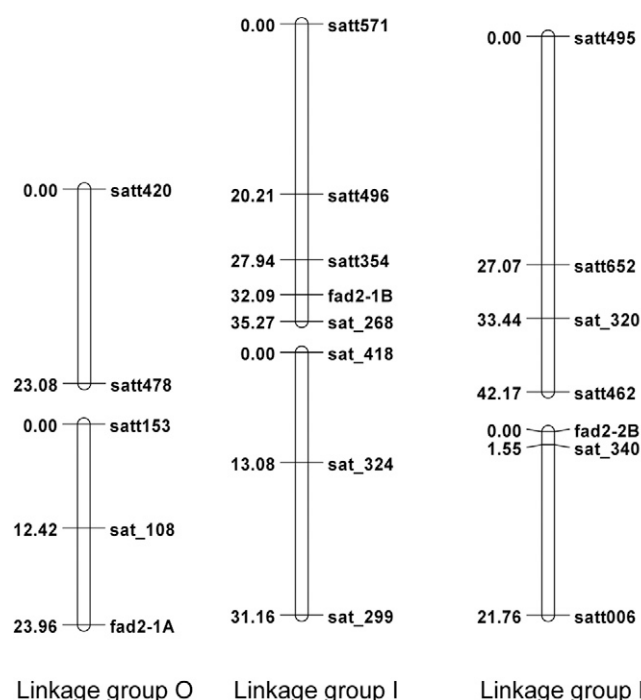


Figure 2. Genetic map of Linkage Groups I, O, and L for the FAF population (N97-3363-3 × PI423893) constructed with JoinMap 3.0. The estimated locations of *FAD2-1A*, *FAD2-2A*, and *FAD2-2B* are presented. It should be noted that mapping of *FAD2-2B* on Linkage Group L in the FAF population is not in agreement with linkage analysis in the FAS population. The *FAD2-2B* isoform probably maps within the interval of the simple sequence repeat markers sat_340 and satt006 on Linkage Group L, as indicated by the linkage map of FAS population (N98-4445A × PI423893). The *FAD2-2A* isoform also colocalizes with *FAD2-2B* on Linkage Group L. Genetic distances were estimated with Kosambi's mapping function. Linkage groups were drawn on an equal scale.

patterns of the CAPS markers for *FAD2-1B* and *FAD2-2B*, as well as of the ASPE and SBE assays for *FAD2-1A*, were utilized for linkage analysis and association tests with fatty acid content in the FAF and FAS populations.

Mapping the Isoforms of the *FAD2-1* and *FAD2-2* Genes

Linkage analysis was conducted for 147 polymorphic SSR markers genotyped in the FAF population, which cover the 20 linkage groups of the soybean genome, as well as the CAPS, ASPE, and SBE assays with both JoinMap 3.0 (Van Ooijen and Voorrips, 2001) and MAPMAKER/EXP 3.0 (Lander et al., 1987). Since, for the linkage analysis performed with MAPMAKER/EXP 3.0, the recombination frequency among markers and the map distances were slightly biased by analyzing the FAF F_5 -derived population as a recombinant inbred line population, only the linkage analysis performed with JoinMap 3.0 is presented here (Fig. 2). It should be noted that linkage could not be declared among all markers on the same linkage group, probably due to the relatively small size of the FAF population. The

FAS population was used to validate the localization of the *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* isoforms.

Mapping of *FAD2-1A* at the end of Linkage Group O near the SSR marker sat_108, and of *FAD2-1B* in the interval of sat_268 and satt354 on Linkage Group I, for both FAF and FAS populations, coincide with the locations proposed by Schlueter et al. (2007). The congruency in the localization of *FAD2-1A* and *FAD2-1B* verifies the efficiency of the developed assays and the efficacy of linkage analysis as a mapping tool. However, a discrepancy in the linkage analysis of FAF and FAS populations was revealed on Linkage Group L, both with JoinMap 3.0 and MAPMAKER/EXP 3.0. The *FAD2-2A* and *FAD2-2B* isoforms mapped in the interval of the SSR markers sat_340 and satt462 for the FAF population, while for the FAS population they were localized in the interval of sat_340 and satt006. Considering the difference in FAF and FAS population sizes, which implies the limited mapping precision of the FAF population, and the location proposed by Schlueter et al. (2007), *FAD2-2A* and *FAD2-2B* probably map within the interval of sat_340 and satt006.

Associations between Molecular Markers and Fatty Acid Content or Maturity Date

Overview of Cosegregation Analysis

Segregation of the *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* allele-specific markers, as well as SSR markers in the vicinity of the isoforms on Linkage Groups O, I, and L, was used to test for association with the content of unsaturated fatty acids in the oil and with maturity date (Table 1). Fatty acids and maturity date BLUPs were obtained for each line from the analysis of each environment separately and from the combined analysis across all environments using mixed models in SAS 9.1. Best linear unbiased predictors for each trait were used in single-factor analysis with *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* allele-specific and SSR marker classes as the independent variables. Although maturity effects were significant for the majority of fatty acid traits and environments tested in the FAE, FAF, and FAS populations, they were not accounted in the computation of BLUPs. Instead, single-factor analysis was also performed with maturity date included as a covariate to determine whether the associations between fatty acid content and molecular markers were affected by maturity. All markers that were significantly associated ($P < 0.05$) with the unsaturated fatty acid content and with maturity date were included in Tables 2, 3, and 4.

Cosegregation Analysis in the Proximity of *FAD2-1A*, *FAD2-1B*, and *FAD2-2B*

On Linkage Group I, segregation of the *FAD2-1B* CAPS marker was associated with changes in oleate and linoleate content in the FAS population. The *FAD2-1B* CAPS marker explained 3.3 and 3.5% of oleate and linoleate genetic variation, respectively, when maturity date was included as a covariate (Table 3). The genetic

locus coincided with a maturity quantitative trait locus (QTL), which explained 6.0% of the genetic variation for maturity. The high-oleate and low-linoleate alleles were inherited from PI423893 and the difference between the means of the homozygous genotypic classes was -13.6 and 11.6 g kg^{-1} , respectively (Table 3). The allele for late maturity was also inherited from PI423893, as expected from the positive correlation between maturity and oleic acid seed content in the FAS population (data not shown). The SSR markers *satt354* and *sat_268*, located 4.91 and 6.07 cM from the *FAD2-1B* isoform, were also supportive of the minor oleate QTL in the proximity of *FAD2-1B* in both FAS and FAE populations (Tables 3 and 4). No significant association was detected on Linkage Group I for the *FAD2-1B* CAPS marker or the flanking SSR markers with fatty acid phenotypes in the FAF population. Thus, the minor oleate QTL in the vicinity of *FAD2-1B* was detected in the FAS and FAE populations but was not validated by the FAF population, which had the smallest population size.

Analysis of the FAE population indicated the existence of a minor oleate QTL in the proximity of the *FAD2-2B* isoform on Linkage Group L. The oleate QTL, which mapped on *sat_340* (13.86 cM from *FAD2-2B* according to the linkage map of the FAS population) explained 2.2% of oleate genotypic variation and was detected in all six environments tested for the FAE population (data not shown). The high-oleate allele at this locus was inherited from the N98-4445A parental line and the means of the homozygous genotypic classes differed by 18.8 g kg^{-1} (Table 4). The minor oleate QTL on Linkage Group L was not detected in the FAF or FAS populations, possibly due to their smaller population size, which impacts the power of QTL detection and the magnitude of the QTL effect (Beavis, 1994). Quantitative trait loci with minor effects for oleic, linoleic, and linolenic acid traits have been reported on Linkage Group L at the genomic region where the *FAD2-2A* and *FAD2-2B* isoforms map (Hyten et al., 2004); however, the population used for that study did not segregate for major oleic acid genes, since it was derived from the cross of normal-oleate parental lines, and maturity effects were not taken into consideration.

On Linkage Group O, segregation of the *FAD2-1A* ASPE and SBE markers was not associated with changes in unsaturated fatty acid content in the FAF and FAS populations, which suggests the lack of oleate QTLs in the proximity of *FAD2-1A*.

Cosegregation Analysis for Other Genetic Loci Mapped on Linkage Groups O, I, and L

On Linkage Group O, analysis suggests a minor oleate QTL near the SSR markers *sat_108* and *satt153*, which explained 2.4 and 0.6% of oleate genotypic variation in

Table 2. The means and standard errors (in parentheses) of homozygous genotypic classes for maturity (R8), oleate (18:1), linoleate (18:2), or linolenate (18:3) content, the difference between means, its standard error (in parentheses), and its significance, as well as the R^2 for the difference of the means for molecular markers significantly associated with maturity, oleate, linoleate, and linolenate content in the FAF (N97-3363-3 × PI423893) population.

Trait	Marker (linkage group)	AA [†]	BB [†]	Difference	R^2
		g kg ⁻¹			%
R8 [‡]	<i>sat_340</i> (L)	115.3 (0.5) [§]	113.1 (0.6)	2.2 (0.8)**	6.3
	<i>FAD2-2B</i> (L)	115.6 (0.6)	112.6 (0.6)	3.0 (0.8)***	10.8
	<i>satt006</i> (L)	115.3 (0.5)	112.4 (0.7)	2.8 (0.9)**	9.1
	<i>sat_268</i> (I)	113.6 (0.6)	115.4 (0.7)	-1.8 (0.9)*	3.7
18:1	<i>satt462</i> (L)	416.8 (7.9)	391.6 (9.6)	25.1 (12.4)*	3.8
	<i>satt652</i> (L)	420.5 (8.2)	394.8 (8.0)	25.7 (11.4)*	4.1
18:1/R8 [¶]	<i>satt462</i> (L)	417.2 (7.4)	389.1 (9.1)	28.2 (11.7)*	4.6
	<i>satt652</i> (L)	421.0 (7.5)	391.7 (7.4)	29.3 (10.6)**	5.2
18:2/R8	<i>satt652</i> (L)	402.7 (6.6)	425.8 (6.6)	-23.1 (9.4)*	4.3
18:3	<i>satt462</i> (L)	44.0 (1.4)	49.5 (1.6)	-5.4 (2.1)*	5.8
18:3/R8	<i>satt462</i> (L)	43.6 (1.3)	50.1 (1.6)	-6.4 (2.0)**	7.9

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

[†]AA, lines homozygous for N97-3363-3 allele; BB, lines homozygous for PI423893 allele.

[‡]R8, maturity (days after planting).

[§]Genotypic classes scored as dominant for N97-3363-3 allele are italicized. The mean and standard deviation for the dominant genotypic class corresponds to the pool of homozygous and heterozygous genotypes.

[¶]Maturity used as a covariate during single-factor analysis.

FAS and FAE populations, respectively (Tables 3 and 4). The *satt153* locus coincided with a QTL for maturity explaining 8.9% of maturity genotypic variation in the FAE population. The late-maturity and low-oleate alleles were inherited from N97-3525, as expected by the negative genetic correlation between maturity and oleic acid seed content in the late-maturing FAE population (data not shown). After accounting for maturity effects in the model, however, it was determined that N97-3525 contributed the high-oleate allele (Table 4), which favors the hypothesis of two linked QTLs controlling maturity and oleate content near the SSR marker *satt153*.

Apart from the minor oleate QTL near *FAD2-2B* on Linkage Group L, minor QTLs for oleate content were also mapped on the SSR markers *satt652* and *satt462*. The SSR marker *satt652*, which is located 65.47 cM away from *FAD2-2B* and 19.67 cM from *satt462* according to the linkage map of the FAS population, explained 5.2% of the genotypic variation for the oleate trait in the FAF population (Table 2). The association between oleate and *satt652* was significant at the $P < 0.01$ level in each of the five environments tested for the FAF population (data not shown),

Table 3. The means and standard errors (in parentheses) of homozygous and heterozygous genotypic classes for maturity (R8), oleate (18:1), linoleate (18:2), or linolenate (18:3) content, the difference between means of homozygous genotypic classes, its standard error (in parentheses), and its significance, as well as the R^2 for the difference of the means for molecular markers significantly associated with maturity, oleate, linoleate, and linolenate content in the FAS (N98-4445A × PI423893) population.

Trait	Marker (linkage group)	AA [†]	BB [†]	HH [†]	Difference	R^2
		g kg ⁻¹				%
R8 [‡]	sat_340 (L)	102.3 (0.4)	99.8 (0.4)	100.9 (0.6)	2.5 (0.6)***	7.6
	<i>FAD2-2B</i> (L)	102.2 (0.4) [§]	100.0 (0.4)		2.2 (0.5)***	7.3
	satt462 (L)	102.0 (0.4)	100.5 (0.4)	99.8 (0.8)	1.5 (0.6)**	3.0
	satt652 (L)	102.0 (0.4)	100.5 (0.4)	101.0 (0.6)	1.6 (0.6)**	3.0
	satt354 (I)	100.2 (0.4)	102.0 (0.4)	101.4 (0.6)	-1.8 (0.6)**	4.0
	<i>FAD2-1B</i> (I)	100.1 (0.4)	<u>102.0 (0.4)[¶]</u>		-2.0 (0.5)***	6.0
	sat_268 (I)	100.2 (0.4)	102.3 (0.4)	100.8 (0.6)	-2.1 (0.6)***	5.7
18:1	satt354 (I)	326.1 (3.5)	343.6 (3.8)	338.8 (5.7)	-17.5 (5.2)***	4.9
	<i>FAD2-1B</i> (I)	326.9 (3.4)	<u>342.4 (3.4)</u>		-15.5 (4.7)**	4.6
	sat_268 (I)	328.4 (3.5)	344.3 (3.4)	333.1 (5.6)	-15.9 (5.2)**	3.9
	sat_108 (O)	331.2 (3.7)	344.2 (3.6)	322.1 (5.8)	-12.9 (5.2)* [-31.2 (12.7)] [#]	2.6
18:1/R8 ^{††}	satt354 (I)	327.0 (3.5)	342.6 (3.8)	338.5 (5.7)	-15.6 (5.2)**	3.7
	<i>FAD2-1B</i> (I)	327.8 (3.4)	<u>341.5 (3.3)</u>		-13.6 (4.8)**	3.3
	sat_268 (I)	329.3 (3.5)	343.0 (3.9)	333.4 (5.6)	-13.7 (5.3)*	2.7
	sat_108 (O)	331.2 (3.7)	343.7 (3.6)	323.2 (5.8)	-12.5 (5.2)* [-28.6 (12.7)] [#]	2.4
18:2	satt354 (I)	466.6 (2.9)	453.1 (3.2)	457.8 (4.8)	13.6 (4.3)**	4.3
	<i>FAD2-1B</i> (I)	466.4 (2.8)	<u>454.2 (2.7)</u>		12.2 (3.9)**	4.1
	sat_268 (I)	466.0 (2.9)	451.6 (3.2)	461.9 (4.6)	14.4 (4.3)***	4.6
18:2/R8	satt354 (I)	466.2 (3.0)	453.5 (3.2)	458.0 (4.8)	12.7 (4.4)**	3.6
	<i>FAD2-1B</i> (I)	466.1 (2.9)	<u>454.4 (2.8)</u>		11.6 (4.1)**	3.5
	sat_268 (I)	465.7 (2.9)	452.0 (3.3)	461.8 (4.6)	13.7 (4.5)**	3.9
18:3	sat_340 (L)	55.2 (1.2)	59.1 (1.4)	58.9 (1.8)	-3.9 (1.9)*	1.9
	satt462 (L)	56.3 (1.1)	59.7 (1.3)	53.8 (2.6)	-3.4 (1.7)*	1.7
	satt652 (L)	55.4 (1.3)	59.8 (1.2)	55.3 (1.8)	-4.3 (1.8)*	2.4
18:3/R8	satt652 (L)	55.9 (1.3)	59.5 (1.2)	55.3 (1.8)	-3.6 (1.8)*	1.6

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

[†]AA, lines homozygous for N98-4445A allele; BB, lines homozygous for PI423893 allele; HH, lines heterozygous for N98-4445A and PI423893 alleles.

[‡]R8, maturity (days after planting).

[§]Genotypic classes scored as dominant for N98-4445A allele are italicized. The mean and standard deviation for the dominant genotypic class corresponds to the pool of homozygous and heterozygous genotypes.

[¶]Genotypic classes scored as dominant for PI423893 allele are underlined. The mean and standard deviation for the dominant genotypic class corresponds to the pool of homozygous and heterozygous genotypes.

[#]Dominance effects estimated as the difference between means of the heterozygous genotypic class from the average of the two homozygous genotypic classes, its standard error, as well as its significance (in square brackets).

^{††}Maturity used as a covariate during single-factor analysis.

as well as in the combined analysis across all environments (Table 2). The oleate QTL near satt652 on Linkage Group L was further confirmed with the FAE population, where satt652 explained 0.5% of the genetic variation for the oleate

trait (Table 4). The significant association between satt652 and oleate content was consistent across two of the six environments when maturity was used as a covariate (data not shown). The high-oleate allele of the detected QTL was inherited from the N97-3363-3 and N98-4445A alleles in the FAF and FAE populations, respectively. In addition, minor QTLs for linoleate and linolenate traits were detected near satt652 after accounting for maturity effects. These results confirm other oleate QTL mapping studies currently in progress (H.R. Boerma, personal communication, 2007). In the FAS population, satt652 was not associated with oleate content but explained 1.6% of linolenate genetic variation (Table 3).

Also, on Linkage Group L, the SSR marker satt462 explained 4.6% of the genotypic variation for the oleate trait in the FAF population (Table 2). The segregation of satt462 was significantly associated with oleate content at the $P < 0.05$ level in four of the five environments tested (data not shown), as well as in the overall environment analysis (Table 2). The FAE population provided additional evidence for an oleate QTL in the proximity of satt462. When maturity effects were taken into account, satt462 explained 1.2% of the genetic variation for oleic acid content in the seed oil (Table 4). The associations declared for the satt462 marker and oleate content in the FAE population were consistent across all six environments tested when maturity was used as covariate (data not shown). The high-oleate alleles of the QTL near satt462 on Linkage Group L were inherited from the N97-3363-3 and the N98-4445A parental lines in the FAF and FAE populations, respectively. The FAS population did not confirm the association of oleic acid content with the SSR markers sat_340 and satt462. In the FAF population, satt462 explained 7.9% of the genotypic variation for linolenate content (Table 2), but the QTL conditioning linolenate content could not be validated in the FAE or FAS populations.

CONCLUSIONS

Oleate QTLs with minor effects were detected in the proximity of the *FAD2-1B* and possibly *FAD2-2B* oleate desaturase genes. Thus, the results of this study suggest that no major oleate QTLs cosegregate with the *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* isoforms of the genes encoding the ω -6 desaturase enzymes. Therefore, the *FAD2-1A*, *FAD2-1B*, and *FAD2-2B* isoforms do not appear to encode for the major oleate genes that contribute to the elevated oleic

acid content of the N98-4445A and N97-3363-3 lines. Further research on the *FAD2-2C* isoform as well as other candidate genes implicated in oleate biosynthesis, such as the plastidial desaturase enzymes, is critical to substantiate the results presented here and shed light on the genetic factors that control oleic acid content in soybean oilseeds.

Acknowledgments

We would like to thank William Novitzky at the USDA Nitrogen Fixation Laboratory, Raleigh, NC, for providing training and equipment for the fatty acid analysis; Perry Cregan and Song Qijian at the USDA Soybean Genomics and Improvement Laboratory, Beltsville, MD, for conducting the SBE assay; as well as Carol Griffin for her technical expertise. This project was funded by the USDA Cooperative State Research, Education and Extension Service National Research Initiative, Plant Genome Program, and by the United Soybean Board.

References

- Beavis, W.D. 1994. The power and deceit of QTL experiments: Lessons from comparative QTL studies. p. 250-266. In D.B. Wilkinson (ed.) *Annu. Corn and Sorghum Res. Conf.*, 49th, Chicago, IL. Am. Seed Trade Assoc., Alexandria, VA.
- Benson, D.A., I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, B.A. Rapp, and D.L. Wheeler. 2002. GenBank. *Nucleic Acids Res.* 30:17-20.
- Brim, C.A. 1966. A modified pedigree method of selection in soybean. *Crop Sci.* 6:220.
- Buhr, T., S. Sato, F. Ebrahim, A.Q. Xing, Y. Zhou, M. Mathiesen, B. Schweiger, A. Kinney, P. Staswick, and T. Clemente. 2002. Ribozyme termination of RNA transcripts down-regulate seed fatty acid genes in transgenic soybean. *Plant J.* 30:155-163.
- Burkey, K.O., F.L. Booker, W.A. Pursley, and A.S. Heagle. 2007. Elevated carbon dioxide and ozone effects on peanut: II. Seed yield and quality. *Crop Sci.* 47:1488-1497.
- Burton, J.W., R.F. Wilson, and C.A. Brim. 1983. Recurrent selection in soybeans: IV. Selection for increased oleic acid percentage in seed oil. *Crop Sci.* 23:744-747.
- Burton, J.W., R.F. Wilson, C.A. Brim, and R.W. Rinne. 1989. Registration of soybean germplasm lines with modified fatty acid composition of seed oil. *Crop Sci.* 29:1583.
- Burton, J.W., R.F. Wilson, G.J. Rebetzke, and V.R. Pantalone. 2006. Registration of N98-4445A mid-oleic soybean germplasm line. *Crop Sci.* 46:1010-1012.
- Cardinal, A.J. 2008. Molecular genetics and breeding for fatty acid manipulation in soybean. *Plant Breed. Rev.* 30:259-294.
- Cardinal, A.J., J.W. Burton, A.M. Camacho-Roger, J.H. Yang, R.F. Wilson, and R.E. Dewey. 2007. Molecular analysis of soybean lines with low palmitic acid content in the seed oil. *Crop Sci.* 47:304-310.
- Cardinal, A.J., M. Lee, N. Sharopova, W.L. Woodman-Clikeman, and M.J. Long. 2001. Genetic mapping and analysis of quantitative trait loci for resistance to stalk tunneling by the European corn borer in maize. *Crop Sci.* 41:835-845.
- Chen, J.W., M.A. Iannone, M.S. Li, J.D. Taylor, P. Rivers, A.J. Nelsen, K.A. Slentz-Kesler, A. Roses, and M.P. Weiner. 2000. A microsphere-based assay for multiplexed single nucleotide polymorphism analysis using single base chain extension. *Genome Res.* 10:549-557.
- Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, L. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. *Crop Sci.* 39:1464-1490.
- Fehr, W.R., and C.E. Caviness. 1977. Stages of soybean development. *Spec. Rep.* 80. Agric. and Home Econ. Exp. Stn., Iowa State Univ., Ames.
- Heppard, E.P., A.J. Kinney, K.L. Stecca, and G. Miao. 1996. Developmental and growth temperature regulation of two different microsomal omega-6 desaturase genes in soybeans. *Plant Physiol.* 110:311-319.
- Hyten, D.L., V.R. Pantalone, A.M. Saxton, M.E. Schmidt, and

Table 4. The means and standard errors (in parentheses) of homozygous genotypic classes for maturity (R8), oleate (18:1), linoleate (18:2), or linolenate (18:3) content, the difference between means, its standard error (in parentheses) and its significance, as well as the R^2 for the difference of the means for molecular markers significantly associated with maturity, oleate, linoleate, and linolenate traits in the FAE (N98-4445A × N97-3525) population.

Trait	Marker (linkage group)	g kg ⁻¹		Difference	R^2
		AA [†]	BB [†]		
R8 [‡]	satt153 (O)	126.2 (0.6)	133.3 (0.6)	-7.1 (0.9)***	8.9
	satt652 (L)	129.4 (0.6)	131.4 (0.8)	-2.0 (1.0)*	0.6
18:1	sat_340 (L)	438.1 (4.4)	425.8 (4.3)	12.3 (6.1)*	0.9
	satt652 (L)	439.6 (3.0)	423.1 (4.6)	16.5 (5.5)**	1.4
	satt153 (O)	441.8 (3.6)	427.7 (3.7)	14.1 (5.1)**	1.2
	satt354 (I)	429.3 (3.8)	441.6 (3.9)	-12.3 (5.5)*	0.9
	satt462 (L)	439.7 (4.0)	427.7 (3.8)	12.0 (5.5)*	0.8
18:1/R8 [§]	sat_340 (L)	441.5 (3.5)	422.7 (3.4)	18.8 (4.9)***	2.2
	satt462 (L)	440.4 (3.2)	425.8 (3.0)	14.6 (4.4)**	1.2
	satt652 (L)	438.8 (2.4)	428.4 (3.6)	9.9 (4.4)*	0.5
	satt153 (O)	428.9 (3.0)	439.1 (3.0)	-10.2 (4.3)*	0.6
	satt354 (I)	430.5 (3.0)	440.4 (3.1)	-9.9 (4.3)*	0.6
18:2	satt462 (L)	417.8 (3.7)	428.7 (3.5)	-10.9 (5.1)*	0.8
	sat_340 (L)	417.9 (4.0)	431.0 (3.9)	-13.1 (5.6)*	1.3
	satt652 (L)	418.5 (2.8)	432.9 (4.2)	-14.3 (5.1)**	1.2
	satt153 (O)	416.9 (3.3)	429.1 (3.4)	-12.1 (4.7)**	1.0
18:2/R8	satt462 (L)	417.3 (3.1)	430.4 (2.9)	-13.1 (4.2)**	1.2
	sat_340 (L)	414.9 (3.3)	433.8 (3.2)	-18.8 (4.6)***	2.6
	satt652 (L)	419.8 (2.3)	428.4 (3.5)	-8.6 (4.2)*	0.4
	satt153 (O)	428.1 (2.8)	419.4 (2.9)	8.7 (4.1)*	0.5
18:3	satt153 (O)	32.4 (0.2)	33.2 (0.2)	-0.9 (0.3)**	1.6
	satt354 (I)	33.3 (0.2)	32.2 (0.2)	1.1 (0.3)***	2.4
	sat_268 (I)	33.1 (0.2)	32.5 (0.2)	0.6 (0.3)*	0.9
18:3/R8	satt354 (I)	33.2 (0.2)	32.3 (0.2)	0.9 (0.2)***	1.9
	satt462 (L)	32.6 (0.2)	33.2 (0.2)	-0.6 (0.3)*	0.6
	sat_268 (I)	33.1 (0.2)	32.5 (0.2)	0.6 (0.3)*	0.8
	satt652 (L)	33.1 (0.1)	32.4 (0.2)	0.7 (0.3)**	1.0

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

***Significant at the 0.001 probability level.

[†]AA, lines homozygous for N98-4445A allele; BB, lines homozygous for N97-3525.

[‡]R8, maturity (days after planting)

[§]Maturity used as a covariate during single-factor analysis.

- C.E. Sams. 2004. Molecular mapping and identification of soybean fatty acid modifier quantitative trait loci. *J. Am. Oil Chem. Soc.* 81:1115-1118.
- Keim, P., T. Olson, and R. Shoemaker. 1988. A rapid protocol for isolating soybean DNA. *Soybean Genet. Newsl.* 15:150-152.
- Kinney, A.J. 1995. Improving soybean seed quality, p. 101-113. *In* Induced Mutations and Molecular Techniques for Crop Improvement: Proc. Symp., Vienna. 19-23 June 1995. IAEA, Vienna.
- Kinney, A.J., and S. Knowlton. 1998. Designer oils: The high oleic acid soybean. p. 193-213. *In* S.H.S. Roller (ed.) Genetic modification in the food industry. Chapman and Hall, London.
- Lander, E.S., P. Green, J. Abrahamson, A. Barlow, M.J. Daly, E.S. Lincoln, and L. Newburg. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174-181.
- Miquel, M., and J. Browse. 1992. *Arabidopsis* mutants deficient in polyunsaturated fatty acid synthesis: Biochemical and genetic characterization of a plant oleoyl-phosphatidylcholine desaturase. *J. Biol. Chem.* 267:1502-1509.
- Ng, P.C., and S. Henikoff. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31:3812-3814.
- Ohlrogge, J., and J. Browse. 1995. Lipid biosynthesis. *Plant Cell* 7:957-970.
- Okuley, J., J. Lightner, K. Feldmann, N. Yadav, E. Lark, and J. Browse. 1994. *Arabidopsis* FAD2 gene encodes the enzyme that is essential for polyunsaturated lipid synthesis. *Plant Cell* 6:147-158.
- Oliva, M.L., J.G. Shannon, D.A. Sleper, M.R. Ellersieck, A.J. Cardinal, R.L. Paris, and J.D. Lee. 2006. Stability of fatty acid profile in soybean genotypes with modified seed oil composition. *Crop Sci.* 46:2069-2075.
- SAS Institute. 2004. SAS/STAT software. Release 9.1.3. SAS Inst., Cary, NC.
- Schlueter, J.A., I.F. Vasylenko-Sanders, S. Deshpande, J. Yi, M. Siegfried, B.A. Roe, S.D. Schlueter, B.E. Scheffler, and R.C. Shoemaker. 2007. The FAD2 gene family of soybean: Insights into the structural and functional divergence of a paleopolyploid genome. *Crop Sci.* 47:S-14-S-26.
- Shoemaker, R., P. Keim, L. Vodkin, E. Retzel, S.W. Clifton, R. Waterston, et al. 2002. A compilation of soybean ESTs: Generation and analysis. *Genome* 45:329-338.
- Song, Q.J., L.F. Marek, R.C. Shoemaker, K.G. Lark, V.C. Concibido, X. Delannay, J.E. Specht, and P.B. Cregan. 2004. A new integrated genetic linkage map of the soybean. *Theor. Appl. Genet.* 109:122-128.
- Tang, G.Q., W.P. Novitzky, H.C. Griffin, S.C. Huber, and R.E. Dewey. 2005. Oleate desaturase enzymes of soybean: Evidence of regulation through differential stability and phosphorylation. *Plant J.* 44:433-446.
- Van Ooijen, J.W., and R.E. Voorrips. 2001. Joinmap 3.0: Software for the calculation of genetic linkage maps. *Plant Res. Int., Wageningen, the Netherlands.*
- Wilcox, J.R., and J.F. Cavins. 1992. Normal and low linolenic acid soybean strains: Response to planting date. *Crop Sci.* 32:1248-1251.
- Wilson, R.F. 2004. Seed composition. p. 621-669. *In* H.R. Boerma and J.E. Specht (ed.) Soybeans: Improvement, production, and uses. Agron. Monogr. 16. 3rd ed. ASA, CSSA, and SSSA, Madison, WI.
- Wilson, R.F., J.W. Burton, and C.A. Brim. 1981. Progress in the selection for altered fatty acid composition in soybeans. *Crop Sci.* 21:788-791.
- Wilson, R.F., J.W. Burton, V.R. Pantalone, and R.E. Dewey. 2002. New gene combinations governing saturated and unsaturated fatty acid composition in soybean. p. 95-113. *In* T.M. Kuo and H.W. Gardner (ed.) Lipid biotechnology. Marcel Dekker, New York.